THE INFANT INDUSTRY ARGUMENT FOR PROTECTION: A REEVALUATION

BY

G. CRIEL*

1 INTRODUCTION

Over the years, many different definitions of an 'infant industry' have been formulated. At this moment, perhaps the most widely accepted definition is the one provided by Westphal. An infant industry, he states, 'is any newly established type of activity for which the economy's existing endowment of skills and human capital does not provide immediate technological mastery.'[1] Technological mastery consists of the command of technological knowledge as illustrated in the ability to use it effectively in the production process. Infant industry protection is thus meant to create the necessary incentives to touch off an investment in learning in order to reach technological mastery in a particular industry. Not every infant industry, however, needs by definition infant industry protection. In fact, the provision of promotional incentives will depend on whether there would otherwise be sufficient incentives to begin the investment in knowledge acquisition.

It is the purpose of this article to determine the general cases in which infant industry protection is justified and to formulate a number of recommendations for the implementation of infant industry promotion. On the basis of an evaluation of the 'traditional' and 'modern' approaches to the infant industry argument, the article will conclude that specific recommendations for infant industry promotion can only be made for specific case studies, but that a successful elaboration of specific infant industry policies will depend on respect for a number of basic principles and criteria which, in the author's view, govern the infant industry argument for protection.

2 THE TRADITIONAL INFANT INDUSTRY ARGUMENT

2.1 Definition

The infant industry argument for protection is but one particular case in the general theory of distortions and is therefore based on the existence of a divergence between social and private returns on a particular investment in

* Generale Bankmaatschappij, Brussels, Belgium.
1 Westphal (1981), p. 2.

learning. For the infant industry to be valid, though, free trade theorists emphasize that a number of rather restrictive conditions must be fulfilled in addition to the basic condition of a deviation between social and private returns. These additional conditions have been articulated progressively throughout the writings of the many economists interested in the issues of international trade and protection.

The earliest elaborate formulations of the traditional infant industry argument by Smith, Ricardo and Mill merely stress the fact that production in the protected infant industry should create those skills and experience necessary to bring domestic production costs below foreign production costs. Mill has encapsuled this early version of the infant industry argument in the following statement[2]:

> 'The only case in which, on mere principles of political economy, protecting duties can be defensible, is when they are imposed temporarily in hopes of naturalizing a foreign industry ... A country which has (the necessary) skill and experience yet to acquire, may in other respects be better adapted to the production than those which were earlier in the field ... A protecting duty, continued for a reasonable time, might sometimes be the least inconvenient mode in which the country can tax itself for the support of such an experiment. But it is essential that the protection should be confined to cases in which there is ground for assurance that the industry which it fosters will after a time be able to dispense with it.'

In short, Mill, in agreement with the economists of his time, argues that protection should be confined to those cases in which it is reasonable to expect that the industry, after a learning period, will be able to survive without government intervention. Mill's test for the applicability of the infant industry argument is thus the future private profitability of the protected industry. Bastable, however, remarks that the mere future private profitability of the industry is an insufficient reason for the imposition of a temporary tariff barrier. An additional condition has to be fulfilled, namely that the social cost of protection in the early years be smaller than the social benefits from lower costs of production in the later years, properly discounted. This condition is usually referred to as 'Bastable's test.' Kemp, in a review of the 'Mill-Bastable infant industry dogma,' introduces a third test. He stresses that, even if both Mill's and Bastable's conditions for the validity of the infant industry argument are fulfilled, protection will only be justified if the production process in the considered industry creates dynamic economies of learning *external* to the firm. In other words, production in the particular industry should create nonappropriable (by the innovator) knowledge and experience which can be

2   Mill (1871), p. 92.

tapped at no costs by any new or established producer on the market. If economies of learning are internal to the firm, *i.e.* if firms can only learn from their own experience, noninnovative firms cannot benefit from the knowledge creation of a pioneer firm. In this case, infant industry protection will not be required since the prospect of later profits will induce the companies to start production and absorb initial losses.

The discussion by Kemp sanctioned the classical view, as it exists today, and as it has been espoused by numerous economists such as Marshall, Meade and Johnson, that, in order to make a valid case of promoting an investment in learning, a number of conditions must be met:

(a) there should be a proven wide discrepancy between the private and the social rates of return on the investment in learning (the latter being superior to the former);

(b) this discrepancy should be the consequence of learning economies related to the production process external to the firm but internal to the industry (in the opposite case, there would be no reason to discriminate against one industry by protecting another);

(c) the protected industry should be expected to be privately profitable in the future; and

(d) the social costs of protection in the early years should be smaller than the social benefits from lower costs of production in the later years, properly discounted.

2.2 *The Insufficiencies of the Traditional Infant Industry Argument*

The major merit of the traditional approach to the infant industry argument for protection is that it allows economists to analyze the argument in the sterile environment of a laboratory. Certain basic assumptions are made which permit the classical theorists to dissect and clarify the 'pure' infant industry argument. Such an approach has theoretical and practical value. It permits a better understanding of the situations which require infant industry protection and it puts on guard anyone who is inclined to promote the abusive use of the argument.

However, the assumptions underlying the traditional infant industry argument are not always realistic. International trade in the modern world does not reflect the assumptions that support the Ricardian theory of comparative advantage. By failing to take into consideration the distortions and imperfections which affect the learning process, the traditional approach, despite its fundamental theoretical validity based on the principle of comparative advantage, has reduced the infant industry argument to the status of a 'theoretical curiosity'[3] and has contributed to its disrepute. Indeed, in the classical literature, the field of application of the infant industry argument has been limited to a number of minor deviations from a static optimum. The conditions of its

3   Myint (1964), p. 175.

use have been made restrictive to the point of making it prohibitive. The exponents of the infant industry argument, on the contrary, have meant their theory to be applied to the dynamic conditions of their times. In fact, the argument can be restored to its full value only if it is seen in the overall context of the broader structural and dynamic problems of economic growth and development.

3 THE MODERN ARGUMENTS FOR INFANT INDUSTRY PROTECTION

It is sometimes suggested that Carey, List and Hamilton formulated the first modern, dynamic versions of the infant industry argument. In fact, their approach was intrinsically classic, albeit integrated in a more global framework of commercial policy. Johnson and Graham made the first real attempts to incorporate the argument in a global theory of trade. Manoilesco went some steps further and developed a comprehensive theory of protection and trade, which laid the foundations for some original approaches to infant industry. It is, however, in the 1950's and the early 1960's with the growing interest in the economic problems of the developing countries, that the study of the argument was most intense. References to the argument can be found in the writings of Hirschman, H.G. Johnson, Lewis, Linder, Myrdal, Prebisch, Nurkse, Singer, and more recently, Balassa, Bhagwati and Corden, to cite only a few of the obvious names.

   This section does not intend to provide a comprehensive evaluation of the leading growth and development theories, but mentions these theories only to the extent required by the discussion of the infant industry argument. The overview may therefore not give full credit to the ideas and contributions of the various economic writers referred to in the article.

3.1 The Dual Economy Approach to Infant Industry Protection

3.1.1 Manoilesco's General Theory of Protectionism
At the basis of Manoilesco's theory lies the Ricardian notion that public intervention in the allocation of resources is justified and required if it is intended to correct a wide deviation between the social and the private profitability of a particular investment, or, as Manoilesco extrapolates, of large sectors of the economy as a whole. However, rather than reaching the classical conclusion that these divergencies permit only certain limited 'cases for protectionism,' Manoilesco pleads for a rigid and comprehensive system of protectionism. For Manoilesco, protection becomes the rule rather than the exception and applies not only to 'infant' industries but also to 'semi-adult' and 'adult' industries.

   In a developing country characterized by a dual economy, Manoilesco explains, the productivity of labor is higher in the manufacturing sector than in the agricultural sector. This is due to the vast surplus of rural workers whose

marginal productivity is zero or negative (the 'hidden unemployment'). Paying market salaries to workers in agriculture or paying the market price for foodstuffs and other agricultural products thus amounts to a burdensome transfer of resources from the productive industrial sector to the nonproductive agricultural sector. Consequently, the private profitability of industrial production is well below its social profitability. Protection of the industrial sector, concludes the argument, becomes the necessary policy response to correct this distortion, as it promotes the transfer of labor from the rural sector to the urban industries and eliminates the extra cost of subsidizing agriculture. The domestic redistribution of labor has important dynamic benefits, argues Manoilesco, since after the transfer the formerly unproductive agricultural labor will contribute to the more efficient industrial production. At the same time, the productivity of labor in agriculture will be stimulated to rise and the efficiency of its production process will increase.

3.1.2 Modern Versions of the Manoilesco Type of Protectionism
The basic idea of Manoilesco's argument for protection has been used by several economic writers interested in economic development. Some authors, such as Lewis, Nurkse, Hirschman and Ranis and Fei have stressed the zero or relatively low marginal product of surplus labor in agriculture to justify or explain protection of the industrial sphere of a developing economy. In such a case, it is argued, agricultural workers can shift to industry without noticeably reducing agricultural production. The capital-intensive urban sector will raise the productivity of labor above its level in the rural sector with all the beneficial dynamic effects of such. Furthermore, since 'unlimited supplies of labor' can be pulled out of the agricultural sector at a constant low wage rate, the manufacturing sector can be expected to earn high profits. The reinvestment of these profits will set off the cyclical process of economic development. Other authors emphasize the gap between wages in industry and agriculture. Hagen has put forward an interesting variant of this second approach. He notes that, based on empirical evidence, 'in any economy in which per capita income is rising secularly, the output of manufacturing and mining grows secularly relative to agriculture ... As a result of this trend, except in the unreal case of perfect geographic and occupational mobility of labor, wages in manufacturing industry must be higher than in agriculture ... (Because) of this wage disparity, manufacturing industry, having a real comparative advantage, will be undersold by imports when foreign exchanges are in equilibrium. Protection which permits such industry to exist will increase real income in the economy.'[4] The promotion of industry must raise the private profitability of this sector, so as to compensate for the higher wage rates and to ensure that the socially optimal quantity of labor is attracted from the rural sector. Institutional changes, learning processes and technological evolutions inherent in

4 Hagen (1958), pp. 497–498.

the process of economic growth will then eventually eliminate the sectoral discrepancies in the real cost of labor. At that point, argues Hagen, the manufacturing industry will have 'grown up' and protection should gradually be removed.

3.1.3 Evaluation

Manoilesco's theory and the Manoilesco type of approach to development and protection are unsatisfactory on several counts. First, the dual economy arguments for protection place much emphasis on the factor labor, but do not consider the factor of capital. But, as a matter of fact, interest rates on investment capital tend to be much higher in the rural sectors than in the urban sectors of developing economies, so that capital in agriculture is often over-valued. The productivity of labor could be inverse to that of capital, and therefore, since the urban sector already enjoys various advantages in the field of communications, transportation, information and capital provision in general, the transfer of cheap labor to industry might be an unwarranted benefit. However, no readily available global theoretical answer can be given. The merits of each case must be valued empirically. Second, the dual economy approach fails to take into account the social costs of infant industry promotion. The static social costs, in terms of the loss of consumer surplus (in the case of a tariff) on the one hand, and the marginal inefficient production of the 'infant' good on the other hand, often tend to be underestimated. Third, the two types of labor, industrial *versus* agricultural, are generally not homogeneous and therefore not readily interchangeable. Finally, and most importantly, the dual economy approach to infant industry protection is basically and unrealistically static in nature. Indeed, the model formulates the infant industry argument in a context in which the amount of fixed factors is optimized, but not changed.

3.2 *The Balanced Growth Approach to Infant Industry Protection*

3.2.1 Description

In his *An International Economy,* Myrdal writes that 'one of the difficulties of industrial development in underdeveloped countries, and one of the great hindrances to giving real momentum to a development policy, is that internal demand must be built up simultaneously with supply. The unlikelihood, or, anyhow, the exasperating slowness of any self-engendered process of "natural growth" offers a main explanation why policy interventions are called for. Indeed the entire idea of a policy of economic development is to break away from this low-level equilibrium.'[5] The policy interventions favored by Myrdal are well known: 'Import restrictions afford a means of by-passing altogether the process of "natural growth" and creating at once the necessary demand for

5   Myrdal (1956), p. 276.

a particular industry.'[6] But which industries should be promoted and how intense should the protection be? Myrdal and others argue that the 'vicious circle of poverty' cannot be defeated by a timid piecemeal attempt to promote economic development, but that the only solution is to engineer a massive capital injection on an all-or-nothing basis in several economic sectors at the same time; *i.e.* by adopting a balanced growth approach to development.

The exponents of this development strategy, including the 'dependency' and 'structuralist' writers of the 1950's and 1960's, as well as believers in a New International Order as suggested by the report of the Brandt Commission, contend that by simultaneously promoting a large number of industries which are (a) technically interdependent (*i.e.* vertically integrated), (b) horizontally compatible (so as to create a demand for each other's products), and (c) large enough in scale ('minimum sensibile' investment), the economy of a developing country will benefit from the working of three types of dynamic effects, namely, (a) the 'technical complementarities' of the production of capital goods and consumer goods, (b) the 'dynamic economies' of production in the protected industries, and (c) the 'internal economies' of large-scale production.

In this context, the infant industry argument receives a new dimension. The argument is based, not on one but on three types of dynamic processes. The classical version of the argument, it has been noted, relies on the existence of external 'dynamic economies' resulting from the nonappropriability of learning. The authors adhering to the balanced growth theory of development suggest that two other types of dynamic effects, namely the 'technical complementarities' of production and the 'internal economies' of large-scale production, can provide, or at least reinforce the case for infant industry protection.

a. *The Problem of Investment Coordination*
Various activities or industries may be complementary to each other to the extent that the more one activity or industry is expanded, the more profitable it becomes to increase investment in the other. Current market prices therefore do not always reflect future as well as present demand and supply conditions and possible misallocations of scarce resources can result. The following example drawn from Chenery illustrates the problem: consider two related industries, steel and metal-working, where the metal-working industry's demand is met through imports. If an innovation in the steel industry lowers its cost of production so that domestic production becomes economically feasible, the size of the steel plant to be constructed in the country would be underestimated if it were based on present import figures. Indeed, the lower supply price of steel will increase the demand for steel of the metal-working industry. 'In this case, external economies exist ... because coordinated investment decisions would result in simultaneous investment in steel and

6 *Ibid.*

metal-working and a lower cost of supply of metal products.'[7] The introduction
of a tariff or a subsidy would raise the profitability of the investment in steel
above the free market level up to its 'real' social level and attract additional
resources to the industry. Once the desired level of investment has been
achieved, the protective measures should be repealed.

Cordon points out that the problem of investment coordination can only
exist in the case of 'indivisible' investments. If investments are divisible, a
relatively minor expansion in one industry would be sufficient to check the
responsiveness of the other and the outputs of the two related industries would
expand in step with each other. If a large investment is required to increase
production in an industry, the management of the firms in that industry will be
reluctant to launch any comprehensive investment program of which the
private returns are not evident. The argument for protection in the context of
investment coordination externalities is thus based on a lack of information or
on the existence of 'uncertainties': how do interrelated industries react on the
expansionary investments in one of these sectors? In the case of two industries,
coordination and information exchanges seem the obvious solution. However,
in an economy with p products, m firms and n industries, the first-best solution
of coordinating intersectoral activities becomes difficult to implement. The
second-best policy is to subsidize temporarily the industries with technical
complementarities and thereby induce them to increase production beyond
the level justified by the private money costs.

The general argument for temporary protection based on the absence of a
comprehensive investment coordination is not as strong as it appears, for
several reasons. First, the first-best policy of investment coordination is not an
impossible goal, as illustrated by the inter- and intraindustry consultations
institutionalized in West Germany and Japan. Consultations and information
exchanges need not necessarily include all participants in the economic process
at once, but can be held on a decentralized basis between industries with
strong potential complementarities, thus creating a chain of interrelated dis-
cussions. This point is certainly a valid critique when complementarities are
evident, but is less convincing in the case of a complex integrated development
plan of the type required by the 'big push' version of the balanced growth
theory. Second, protection cannot be a permanent substitute for coordination.
After the economy has stabilized at a higher level equilibrium as a result of the
temporary protection, the issue of coordination is still not resolved. Protection
must therefore not only stimulate the industry with technical complemen-
tarities to produce more than it would without assistance, but it must also
induce investment coordination between the industrial sectors. Finally, as
Corden argues strongly, the argument is weakest in the situation where
uncoordinated industries produce or require traded goods. Indeed, the indus-
try that decides to expand its production for internal or external reasons can

7   Chenery (1959), p. 85.

always opt to import its necessary inputs or export its output. The issue is really pressing only in the case of nontraded goods.

b. *Economies of Scale*
The existence of internal economies of scale is not accepted as a valid justification for infant industry protection in the classical theory of trade. The traditional view is that private companies do have the necessary technical information to predict the cost effects of economies of scale in their industries and will therefore invest in a venture or expand existing production of a good if the foreseeable scale effects are strong enough. Prebisch, on the other hand, suggests that the economies of scale should be seen as a 'logical extension of the infant industry argument,'[8] meaning that the productivity of an investment in technological mastery does not only depend on time but also on the size of the market. He believes that even a market including all developing countries would not be large enough to push down the domestic production of an 'infant' commodity below the world price, and he therefore pleads for a generalized system of preferences for the industrial products of the developing countries on the industrialized countries' markets. Prebisch does not elaborate on the nature of the 'economies of scale' the domestic industry could benefit from, nor does he propose to narrow down the field of application of his argument to a number of selected industries with high potential scale effects. He argues, more implicitly than explicitly, that at a higher scale of production the industrial learning process is qualitatively different and quantitatively more important than at a lower scale of production.

Corden agrees with Prebisch that the precise amount and time-span of protection, justified by dynamic economies of time, may depend on the static economies of scale effects. He mentions various possible relationships between scale and learning: (a) learning per unit of output may increase simultaneously with both scale and time, and hence rise with cumulative output; (b) the productivity of learning may only be profitable if it can be amortized over a large production output, and therefore the investment will only be undertaken if economies of scale are anticipated, and (c) the nature of the learning process may be different at different scales of production because of the differences in techniques and processes. On the whole, however, Corden concludes, these considerations do not alter the fact that static economies of scale, taken alone, are not a valid justification for infant industry protection.

Havrylyshyn and Wolf support Corden's view that economies of scale can be an efficient catalyst in the creation of knowledge in developing countries, but suggest that protection (or subsidization) might not necessarily be the best means to induce high-scale domestic production and promote high learning rates in an infant industry. In a world where developing countries trade only with developed countries, the argument runs, more sophisticated and capital-

8   Prebisch (1964), p. 80.

intensive goods can initially only be produced in a developing country if the domestic market for the industry is cut off from foreign competition by costly tariff barriers. The huge social welfare losses associated with protection are well known. The developing countries, however, Havrylyshyn and Wolf reason, can take advantage of the dynamic benefits of trade among themselves to reduce the cost of infant industry promotion. The existence of other developing countries with less efficient production processes in the supply of the particular 'infant' product or with no domestic production of that commodity, provides a natural comparative advantage (*e.g.* cultural elements, transportation costs, marketing, . . .) to the infant exporters of the developing country. As trade proceeds, the production of the 'infant' product will become more efficient through the play of learning effects and scale economies and, eventually, exports to developed countries will occur.

3.2.2 Evaluation
The across-the-board import substitution policies adopted by many developing countries in the 1950's and 1960's brought along, in most cases, disastrous results. 'The irony is that the (developing) countries themselves, by protecting industry, prevent their own industrialization. Their import restrictions keep out the capital-intensive heavy manufactures that are necessary to build, and provide materials for, labor-intensive industries. Domestic production of capital-intensive manufactures raises the economy's capital requirements and thus raises the amount of domestic saving required to achieve a given growth rate. But the reduction in national income accompanying protection makes savings scarcer. Only landed countries can industrialize through protection because only they can afford it.'[9] Conceptually, the weaknesses of the undiscriminating balanced growth approach to development have been highlighted by numerous writers, such as Balassa, Bhagwati, Chenery, Lal, and many others. These economists emphasize the importance of eventempered cost-efficiency considerations in the selection of a growth and development strategy. Empirically, the drawbacks of the protectionist approach to development have been confirmed in almost all studies undertaken on the subject. More specifically, two major projects on foreign trade regimes and their effects on economic development have been conducted by the OECD and the NBER, providing empirical evidence and economic argumentation on a number of important trade and development issues. The two research projects stress the high average tariff levels for manufactures in developing countries. The OECD report argues that most of the present industrialized countries (with maybe the exception of Japan) used substantially lower tariffs in the course of their development. Similarly, the NBER country studies suggest that restrictive foreign trade regimes with higher tariffs and high effective exchange rates on imports as compared to effective exchange rates on exports bring

9   Travis (1964), p. 245.

along lower export performances. Lower export performance, again, is empirically associated with poorer growth rates.

But where does all this leave the infant industry argument for protection? The balanced growth approach to development does not deal with the important issue of temporary protection in a number of carefully selected industries or activities, but proposes an across-the-board protection of the developing economies. It is true that by differentiating the intensity of protection among sectors, certain activities can be favored over others. Nevertheless, the exponents of the import substitution approach think more of the 'development' or 'young country' arguments for protection than of the more restrictive 'infant industry' argument. The limitations of the undiscriminating import substitution policies constitute a healthy warning that the arguments for protection in general, and those for infant industry in particular, cannot be used on a generalized basis without threatening the effectiveness of the public adjustment techniques. As Nurkse admits, 'while it is not to be denied that import restriction can help a policy of balanced domestic investment,' it should be used carefully to avoid driving up too far the costs of domestic substitute production, to which he adds that 'import restrictions imposed in spite of such unfavorable effects can be justified only on the ground of future benefits, which is the (traditional) infant industry argument for protection.'[10]

3.3 *The Unbalanced Growth Approach to Infant Industry Protection*

3.3.1. Hirschman's Vertical Linkages

a. *Description*
In his critique of the balanced growth doctrine of development, Fleming argues that in a situation where the supplies of labor, capital and other resources are not perfectly elastic, the simultaneous creation of a wide range of industries is likely to result in substantial external diseconomies to the new industries themselves because of the negative price effects of the high demand for a limited amount of resources. Fleming therefore pleads for the creation of vertical groups of industries, with clearly defined technical complementarities, which he calls 'vertical linkages' between industries at different stages of production. The concept of 'vertical linkages' is based on the assumption that technically given and fixed input-output coefficients exist between different activities in an economy. The expansion of one activity will increase the demand for related inputs produced by other firms through the play of 'backward linkages.' Similarly, the same activity expansion will increase the supply of inputs of downstream sectors ('forward linkages') and through the push effect is expected to stimulate production in these sectors.

This idea of forward and backward linkages lies at the basis of Hirschman's

10   Nurkse (1957), pp. 17–18.

approach to economic development. Hirschman believes that the developing countries would benefit by concentrating their investments in those 'strategic' sectors of their economies where the linkage effects are expected to be strongest. The dynamic secondary effects of an initial investment would then be greatest. The strategic parts of the economy, according to Hirschman, are those where the input-output relationships are thickest, representing industries which buy the largest share of their inputs from other industries, or sell the largest part of their outputs to other industries, or both.

Hirschman stresses that his linkage approach to economic development can only bear results if there is sufficient domestic demand for the commodities produced by the vertically integrated industries. The best indication of the existence of a domestic demand for a particular commodity not produced domestically is, according to Hirschman, the volume of imports of the product. Once the direct and indirect imports of a commodity i reach a certain threshold Ti, determined by the minimum economic size of domestic production, domestic investment in the particular industry will become economically desirable. But this automatic substitution process is not likely to function adequately in the realistic conditions of a developing country. Hirschman therefore favors a limited degree of protectionism to set off the domestic production of the particular commodity, which must accelerate the dynamic import substitution mechanism.

Against this background, Hirschman's infant industry argument for protection can be summarized as follows: temporary protection of an industry in a developing country is economically justified and desirable, provided that (a) the industry is on or near the threshold of minimum domestic production, (b) the industry has strong vertical linkage effects (protectionism 'in depth' rather than 'in breadth'), and (c) protection is not awarded to too many industrial sectors at the same time.

b. *Evaluation*

Hirschman's approach to economic development in general and to infant industry protection in particular has some weaknesses. First, the 'threshold' related to the minimum size of domestic production has to be defined. This assumes (a) that it is technically possible to determine this minimum level, and (b) that once this minimum size is attained, the domestic producer of the import substitute will be able to compete with the foreign supplier. Second, Hirschman's approach assumes that there is already a certain amount of autonomous income in the developing country. Often, however, the development problem in such countries is related to the 'low level equilibrium' of their economies. Third, the linkage approach focuses exclusively on the technically possible links between the inputs and outputs of various industries without considering the economic realization of these linkages, a realization which depends on the relative costs and prices of the inputs and outputs. The introduction of international trade and the principle of comparative advantage

question the economic desirability of the domestic production of certain commodities. Fourth, the theory of industrial linkages depends to a large extent on the responsiveness of private enterprise in developing countries to bottleneck situations in the economy which entail profitable opportunities of increasing investment. Riedel however argues that 'the mere existence of linkages is insufficient to ensure that inducements will indeed be generated.' Such a view presupposes that demand will create its own supply, an 'assumption . . . at this point in history . . . at least a naive as the reverse assumption was for the industrial countries in earlier times.'[11] Finally, Myint challenges the generalization, implicit in Hirschman's theory, that the expansion of the manufacturing sector is likely to generate stronger linkages with the rest of the economy than the expansion of agriculture. An increase in agricultural production is as likely to stimulate the demand for machinery, tools, and fertilizers, and create additional demand for agricultural products, as an expansion of industry is likely to have dynamic linkage effects. The relative importance of the technical complementarities in respectively agriculture and industry depends on local economic conditions and comparative advantages.

3.3.2 Linder's Representative Demand

a. *Description*
Linder has developed a dynamic theory of trade and development, and an innovative argument for infant industry protection, based on his idea of 'representative demand'. Of central importance to Linder's theory is the concept of foreign exchange gap, which he defines as the divergence between the amount of foreign exchange a developing country earns through its exports, assuming 'the pursuance of optimum expenditure policies and commercial policies formulated on the basis of conventional trade theory,'[12] and the amount of foreign exchange, under the same conditions, the country needs to fully utilize the productive factors in the economy. The emergence of such a gap is due, first, to the existence of certain minimum import requirements as a consequence of the limited substitutability of domestic input factors and input imports ('import minimum'). Input imports are those imports which are assumed to be necessary to avoid underutilization of existing resources. Second, there is a natural ceiling on export earnings ('export maximum') determined by the nature and the amount of the country's exports. The 'export maximum' may fall short of the 'import minimum,' giving rise to a structural foreign exchange gap. This gap 'does not imply overabsorption of goods and services in the ordinary sense. No matter how much the absorption is reduced, the deficit cannot be removed without substituting internal disequilibrium for

11   Riedel (1976), p. 322.
12   Linder (1967), p. 43.

external.'[13] Linder thus explains the foreign exchange gap by the structural constraints inherent in the process of economic development.

The fundamental reason for the existence of the 'import minimum' and the 'export maximum,' Linder argues, is the fact that international trade is not as much the result of differences in relative resource endowments as of the structure of the domestic demand. The more the internal demand of a country (and therefore the domestic production) is representative of the demand in the export markets, the better the chances of increasing the 'export maximum,' because the more likely that domestic learning and returns to scale will be beneficial to the country's exports. As Linder explains, owing to the lack of similarity between the foreign and the domestic demand, 'the developing countries ... cannot export (to the industrial countries) those manufactures they are most efficient at producing. Generally speaking, they are reduced to trying to export manufactures with which they have no experience.'[14] An important step toward increasing the 'export maximum' and limiting the foreign exchange gap of a developing country is the creation in the developing country itself of those industries which have already, or are likely to have in the future, a 'representative demand.' In this context, Linder pleads for trade restrictions and a policy of import substitution which promotes such a domestic production structure. However, import substitution may not be entirely successful because of the need to finance the foreign exchange component of the investments required for the production of import substitution goods. Linder therefore argues that a particular import substitution project will be efficient, and will deserve infant industry protection, if it is a net saver of foreign exchange over its lifetime:

> 'In the infant industry case for protection suggested here, it is ... required that the foreign exchange cost on the input side of the project be more than offset on the output side. This efficiency criterion is weaker, as the project is not required to become internationally competitive ... The question is not one of gaining through the superior allocation of resources: it is one of instituting certain activities behind trade controls which, through foreign exchange savings and increased input imports, enable the country to exert a leverage on capacity growth.'[15]

Linder even broadens his argument by stating that an investment will also deserve infant industry protection, even if it is expected to remain inefficient during its lifetime, provided that foreign exchange savings will be made during the life of subsequent projects of the same type.

13 *Ibid.*, p. 51.
14 *Ibid.*, p. 37.
15 *Ibid.*, pp. 93–94.

b. *Evaluation*
The basic idea of Linder's approach to trade and protection fits into the theory of unbalanced growth and development. The developing countries, Linder suggests, should concentrate their domestic production efforts on a number of carefully selected sectors as determined by the representativeness of the demand and the ability to produce foreign exchange. In doing so, he acknowledges that static effects are not the only positive effects of trade and are probably the least important.

The major limitation of Linder's approach, which in fact is also the major weakness of all suggested substitution policies in the context of the unbalanced growth approach to development, is that (once again) the cost-effectiveness of the substitution policies is neglected. Even though Linder, Hirschman and others provide 'scientific' yardsticks for the selection of sectors to be protected, the recommended trade policies are not fundamentally different from those suggested by the exponents of the other import substitution theories.

Finally, Linder's argument that international trade flows will radiate toward proximate and similar markets is contradicted by the pull comparative advantages are expected to have on exchange patterns. Linder uses his theory of the 'representativeness of the demand' to explain that, *ceteris paribus,* developing countries will mainly trade with each other. It is surprising that Linder's approach, which seems theoretically incorrect, has nevertheless given solid predictions. The real reasons for this might be: (a) the importance of transportation costs and geographical proximity, (b) the intra-industry specialization of developing countries based on economies of scale, (c) the reliance of inward-looking economies on demand factors of other economies for their exports, and (d) a policy of 'collective self-reliance' based on the political desire to 'delink' their economies from those of the developed countries.

4 EVALUATION AND CONCLUSIONS

In the above discussion of the traditional and modern theories of infant industry protection the various strengths and weaknesses of the different approaches to the argument have been evaluated. However, none of the existing formulations of the infant industry argument have been fully convincing. The classical approach to the argument, based on the correct perception that economic production and exchanges are best served by respect for the principle of comparative advantage, is on the whole too static and too restrictive to be of any practical value in real world situations. The modern approaches to infant industry protection, even though some are basically dynamic in nature, fail to provide a convincing theoretical background for the argument. These concluding paragraphs will attempt to reformulate the argument for infant industry protection by combining the best of both approaches, namely, on the one hand, the basic principle of comparative advantage of the classical theory of trade, and on the other hand, the dynamic approach and the

responsiveness to specific market situations of the modern theories.

It does not seem possible nor desirable to provide the policymakers in the developing countries with ready-made answers to the fundamental questions of how and when to implement infant industry promotion. The merits and characteristics of specific cases must be valued empirically. However, a precise and potentially successful elaboration of specific infant industry policies will depend to a large extent on the understanding of the basic principles which govern the economic viability of promotional incentives in general and infant industry protection in particular. The following paragraphs propose to summarize the fundamental conditions and principles which, in the author's view, must be respected in the identification and implementation of valid cases for infant industry promotion, without, however, strapping the argument into an ideological straitjacket. This reevaluation of the infant industry argument draws extensively on the evidence gathered and the ideas formulated by the various economic theorists and researchers quoted in the previous sections, and its merit lies mainly, though not exclusively, in the unifying and classifying of presently scattered research and ideas. In doing so, it must be repeated, the argument is based on the fundamental classical principle of comparative advantage but is set in a dynamic context which takes into account the inperfections of the economic sphere in the developing countries.

The governing principles of this revised view on infant industry protection are summarized in the propositions below.

(a) Infant industry protection can only be justified and desirable in the case of an investment in the acquisition of technological mastery which is socially desirable but privately unprofitable. The social and private profitability rates of alternative projects in learning must be determined coherently and comprehensively by modern cost-benefit techniques so as to select the socially most desirable ventures.

(b) The learning process should be of a dynamic, reciprocal nature, where learning in one firm stimulates learning in another firm, and *vice versa,* so as to touch off a self-sustaining process of knowledge creation within an industry. Even though temporary protection might be justified in the case of learning externalities or imperfections limited to one firm, it would not be an argument for infant industry protection.

(c) The discrepancy between the social and private profitability rates of a particular investment in learning can be the result of production externalities (where a producer does not have to pay for a production input) and/or of market imperfections or distortions which affect the learning process. The infant industry argument based solely on production externalities is too restrictive and unrealistic. In many cases, markets are imperfect and prices do not reflect the relative scarcity values of various goods needed in an investment in the acquisition of technological mastery.

(d) Infant industry protection should be awarded to those industries or activities which enjoy a comparative advantage over foreign suppliers or which

are likely to do so in the near future. The reliance on the concept of comparative advantage ensures the dynamism of the argument for infant industry promotion, is a good indication of its social desirability, and, to the extent that future comparative advantages are taken into account, makes it forward-oriented.

(e) The social losses associated with the infant industry protection in the early phases of its realization must be recovered in the later stages of production, properly discounted, and this on an industry-wide basis. The social benefits need not be equally distributed over time and over the various firms or investments within the industry. In fact, it is argued here that an investment in learning will be desirable, even if, taken on its own, it does not become socially profitable during its lifetime, provided that it allows for ulterior or simultaneous profitable investments in the same industry. Of importance is that, within the given time limit of the duration of the initial infant industry project, the whole set of related activities becomes socially profitable.

(f) The industry, but not necessarily all firms within the industry, should be able to survive without protection after the initial learning period during which the industry is promoted. Those firms which cannot survive after this date should not be allowed to receive additional protection.

(g) Future and present comparative advantages can be influenced by the catalytic effect of economies of scale and investment coordination efforts on the process of acquisition of technological mastery. Infant industry cannot be justified on the basis of economies of scale or investment coordination taken alone, but since they can accelerate knowledge creation and even improve the quality of knowledge acquisition, they become, to use Prebisch's words, a natural extension and reinforcement of the infant industry argument.

(h) Infant industry protection should be unbiased against export-oriented industries, unless particular dynamic effects are proven to exist in import substitution sectors. Empirical and theoretical evidence seems to indicate that export-oriented macro-economic policies, which provide similar incentives to sales in domestic and foreign markets, lead to resource allocations according to comparative advantages, allow for greater capacity utilization, permit the exploitation of economies of scale, accelerate and improve technological learning, and, in labor surplus countries, contribute to higher employment.

(i) Infant industry protection should be unbiased against agricultural activities unless particular dynamic benefits are credited to the industrial sector. Empirical evidence indicates that industrial investments have more positive dynamic effects than investments in the agricultural sector. On the other hand, market imperfections and distortions might be more disruptive in the agricultural sphere of an economy.

(j) Infant industry protection should be implemented administratively so as to (1) minimize the costs to society (preference of first-best solutions, even at short term losses), (2) ensure the temporality of protection, and (3) rationalize, simplify and generalize the structure of protection.

The conclusion of the article has attempted to unify a considerable body of literature on the issue of trade and infant industry protection into a series of major propositions that constitute a generalized theory of infant industry promotion. Besides the intrinsic value of the summarizing and reevaluating effort, this has permitted a number of insights into the existing theories on infant industry protection.

REFERENCES

Balassa, B., 'Reforming the System of Incentives in Developing Countries,' *World Development,* III (1975), pp. 365–382.

Balassa, B., *Policy Reform in Developing Countries,* Oxford, 1977.

Baldwin, R.E., 'The Case against Infant Industry Tariff Protection,' *Journal of Political Economy,* LXXVII (1969), pp. 295–305.

Bastable, C.F., *The Commerce of Nations,* London, 1921.

Bhagwati, J.N. (ed.), *International Trade: Selected Readings,* Cambridge, 1982.

Chenery, H.B., 'The Interdependence of Investment Decisions,' in: M. Abramovitz, *et al., The Allocation of Economic Resources,* Stanford, 1959.

Chenery, H.B., *Structural Change and Development Policy,* New York, 1979.

Corden, W.M., *Trade Policy and Economic Welfare,* Oxford, 1974.

Fleming, J.M., 'External Economies and the Doctrine of Balanced Growth,' *Economic Journal,* LXV (1955), pp. 241–256.

Graham, F.D., 'Some Aspects of Protection Further Considered,' *Quarterly Journal of Economics,* XXXVII (1923), pp. 199–227.

Grubel, H.G., 'The Anatomy of Classical and Modern Infant Industry Arguments,' *Weltwirtschaftliches Archiv,* XCVII (1966), pp. 325–324.

Haberler, G., 'Some Problems in the Pure Theory of International Trade,' *Economic Journal,* LX (1950), pp. 223–240.

Hagen, E., 'An Economic Justification of Protectionism,' *Quarterly Journal of Economics,* LXXII (1958), pp. 496–514.

Hamilton, A., *Papers on Public Credit, Commerce and Finance,* edited by S. McKee, New York, 1934.

Havrylyshyn, O., and M. Wolf, *Trade Among Developing Countries: Theory, Policy Issues, and Principal Trends,* Washington, DC, 1981.

Hirschman, A.O., *The Strategy of Economic Development,* New Haven, 1958.

Johnson, A., 'Protection and the Formation of Capital,' *Political Science Quarterly,* XXIII (1908), pp. 220–241.

Johnson, H.G., 'A New View of the Infant Industry Argument,' in: I. McDougall and R. Snape (eds.), *Studies in International Economics,* Amsterdam, 1970.

Kemp, M.C., 'The Mill-Bastable Infant Industry Dogma,' *Journal of Political Economy,* LXVIII (1960), pp. 65–67.

Lewis, W.A., 'Economic Development with Unlimited Supply of Labor,' *Manchester School of Economic and Social Studies,* XXII (1954), pp. 139–191.

Linder, S.B., *Trade and Trade Policy for Development,* New York, 1967.

List, F., *The National system of Political Economy,* New York, 1966.

Manoilesco, M., *The Theory of Protection and International Trade,* London, 1931.

Marshall, A., *Money, Credit and Commerce,* London, 1923.

Meade, J.E., *Trade and Welfare,* Oxford, 1955.

Mill, J.S., *The Principles of Political Economy,* London, 1871.

Myint, H., 'Infant Industry Arguments for Assistance to Industries in the Setting of Dynamic Trade Theory,' in: R. Harrod and D. Hague (eds.), *International Trade Theory in a Developing World,* London, 1964.

Myint, H., *The Economics of the Developing Countries,* New York, 1974.

Myrdal, G., *An International Economy,* New York, 1956.

National Bureau of Economic Research (NBER), *Foreign Trade Regimes and Economic Development,* several case studies, New York, 1975–76.

Negishi, T., 'Protection of the Infant Industry and Dynamic Internal Economies,' *Economic Record,* XLIV (1968), pp. 56–76.

Nurkse, R., *The Conflict between Balanced Growth and International Specialization,* Istanbul, 1957.

Organization for Economic Cooperation and Development (OECD), *Industrialization and Trade Policies,* several case studies, Oxford, 1970.

Pennisi, G., 'L'Argomento della Industria Nascente: Un Tentativo di Riformulazione,' *Rivista di Politica Economica,* LIX (1969), pp. 671–715.

Prebisch, R., *Toward a New Commercial Policy for Economic Development,* New York, 1964.

Ranis, G. and J. Fei, 'A Theory of Economic Development,' *American Economic Review,* LI (1961), pp. 533–565.

Riedel, J., 'A Balanced Growth Version of the Linkage Hypothesis: A Comment,' *Quarterly Journal of Economics,* XC (1976), pp. 319–322.

Travis, W.P., *The Theory of Trade and Protection,* Cambridge, 1964.

Westphal, L.E., *Empirical Justification for Infant Industry Protection,* Washington, DC, 1981.

Summary

THE INFANT INDUSTRY ARGUMENT FOR PROTECTION: A REEVALUATION

Despite its importance as an exception to the free trade case, the argument for infant industry protection has only sporadically been subjected to close theoretical scrutiny. This article helps to fill this gap by evaluating the various 'traditional' and 'modern' perceptions of the topic. The paper concludes that none of the existing formulations of the infant industry theory are fully satisfactory. The classical view on the argument seems too static and too restrictive, whereas the modern approaches fail to provide a convincing theoretical framework. The proposed reappraisal of the infant industry idea is based on the classical principle of comparative advantage, but places the argument in the dynamic context of modern theories.